## SEQUENCE LISTING

<110>	RIKEN KABUSHIKI KAISHA DNAFORM	
<120>	Method for utilizing the 5'end of mRNA for cloning and analysis	
<130>	1336(PCT)	
	JP 2002-171851 2002-06-12	
<150> <151>	JP 2002-235294 2002-08-12	
<160>	77	
<170>	PatentIn version 3.1	
<210> <211> <212> <213>	1 74 DNA Artificial	
<220> <223>	First strand cDNA primer	
<220> <221> <222> <223>	(73)(73)	
<220> <221> <222> <223>	misc_feature (74)(74) "n" is any nucleotide	
<400> gagaga	1 agaga aaggateetg ceattteatt acetetttet eegcaceega catagatttt	60
ttttt	ttttt ttvn	74
<210> <211> <212> <213>	60 DNA	
<220> <223>	Upper oligonucleotide GN5	
<220> <221> <222> <223>	misc_feature (56)(60)	
<400> agaga	2 gagac ctcgagtaac tataacggtc ctaaggtagc gacctaggtc cgacgnnnnn	60
<210> <211> <212> <213>	60 DNA	
<220> <223>		

<220> <221> <222> <223>	misc_feature (55)(60) "n" is any nucleotide	
<400> agagaga	3 agac ctcgagtaac tataacggtc ctaaggtagc gacctaggtc cgacnnnnnn	60
<210> <211> <212> <213>	DNA	
<220> <223>	Lower oligonucleotide	
<400> gtcgga	4 ccta ggtcgctacc ttaggaccgt tatagttact cgaggtctct ctct	54
<210> <211> <212> <213>		
<220> <223>	primer	
<400> agagag	5 agac ctcgagtaac tataacggtc ctaaggtagc gacctaggtc cgacg	<b>5</b> 5
<210> <211> <212> <213>	45	
<220> <223>	linker	
<400> tctaga	6 tcag gactetteta tagtgteace taaagtetet etete	45
<210> <211> <212> <213>	7 47 DNA Artificial	
<220> <223>	linker	
<220> <221> <222> <223>		
<400> gagaga	7 gaga ctttaggtga cactatagaa gagteetgat etagann	47
<210> <211> <212>		

<220> <223>	Primer 1 (uni-PCR)	
<400> gagaga	8 gaga ctttaggtga cacta	25
<210> <211> <212> <213>	25	
<220> <223>	Primer 2(MmeI-PCR)	
<400> agagag	9 agac ctcgagtaac tataa	25
<210> <211> <212> <213>	44	
<220> <223>	first strand oligo-dT primer	
<220> <221> <222> <223>	misc_feature (43)(43) "v" is A, C or G	
<220> <221> <222> <223>	misc_feature (44)(44) "n" is any nucleotide	
<400> gagaga	10 ngaga ggatccttct ggagagtttt ttttttttt ttvn	44
<210> <211> <212> <213>	45	
<220> <223>	Oligonucleotide Bg-Gsu-GN5	
<220> <221> <222> <223>	misc_feature (41)(45) "n" is any nucleotide	
<400> agaga	11 gagaa ctaggcttaa taggtgacta gatctggagg nnnnn	45
<210> <211> <212> <213>	45 DNA	
<220> <223>	Oligonucleotide Bg-Gsu-N6	
<220>		

WO 03/106672

```
<221> misc_feature
<222> (40)..(45)
<223> "n" is any nucleotide
 <400> 12
 agagagagaa ctaggcttaa taggtgacta gatctggagn nnnnn
                                                                                                              45
<210> 13
<211> 39
<212> DNA
<213> Artificial
 <220>
<223> Oligonucleotide Bg-Gsu-down
 <400> 13
ctggagatct agtcacctat taagcctagt tctctctct
                                                                                                              39
<210> 14
<211> 47
<212> DNA
<213> Artificial
 <220>
<223> Oligonucleotide Bg-Mme-GN5
<220>
<221> misc_feature
<222> (43)..(47)
<223> "n" is any nucleotide
<220>
<221> misc_feature
<222> (39)..(39)
<223> "r" is G or A
 <400> 14
agagagagaa ctaggcttaa taggtgacta gatcttccra cgnnnnn
                                                                                                              47
<210> 15
<211> 47
<212> DNA
<213> Artificial
<220>
<223> Oligonucleotide Bg-Mme-N6
 <220>
<221> misc_feature
<222> (42)..(47)
<223> "n" is any no
           (42)..(47)
"n" is any nucleotide
<220>
<221> misc_feature
<222> (39)..(39)
<223> "r" is G or A
agagagaga ctaggcttaa taggtgacta gatcttccra cnnnnn
                                                                                                              47
<210> 16
```

## WO 03/106672

<211> <212> <213>	40 DNA Artificial	
<220> <223>	Oligonucleotide Bg-Mme-down	
<2222	misc_feature (3)(3) "y" is C or T	
<400> gtygga	16 gatc tagtcaccta ttaagcctag ttctctctct	40
<210> <211> <212> <213>	47	
<220> <223>	oligonucleotide	
<220> <221> <222> <223>	misc_feature (46)(47) "n" is any nucleotide	
<400> gagaga	17 gaga ctttaggtga cactatagaa gagtcctgag aattcnn	47
<210> <211> <212> <213>	45	
<220> <223>	oligonucleotide	
<400> gaatte	18 etcag gactetteta tagtgteace taaagtetet etete	45
<210> <211> <212> <213>	49	
<220> <223>	linker	
<222>	misc_feature (45)(49) "n" is any nucleotide	
<400> agaga	19 gagag cttagatgag agtgactcga gcctaggtcc aacgnnnnn	49
<210> <211> <212> <213>		

<220> <223>	linker	
<222>	misc_feature (44)(49) "n" is any nucleotide	
<400> agagag	20 agag cttagatgag agtgactcga gcctaggtcc aacnnnnnn	49
<210> <211> <212> <213>	43	
<220> <223>	linker	
<400> gttgga	21 ccta ggctcgagtc actctcatct aagctctctc tct	43
<210> <211> <212> <213>	42	
<220> <223>	second linker	
<400> gaattc	22 tacg cetetegate gaaateeega tetaggetag eg	42
<210> <211> <212> <213>	42	
<220> <223>	second linker	
<400> cttaag	23 atgc ggagagcgtg aatcgagttt aaggctagca tc	42
<210> <211> <212> <213>	25	
<220> <223>	primer	
<400> ttagat	24 gaga gtgactcgag cctag	25
<210> <211> <212> <213>	25	
<220> <223>	primer	
<400>	25	

C	ctacgatcgg aatttgagct aagtg		
<	210> 211> 212> 213>	26 17 DNA Artificial	
	220> 223>	primer	
	400> aggaaa	26 acag ctatgac	17
<	210> 211> 212> 213>	16 DNA	,
	220> 223>	primer	
	(400> staaaa	27 cgac ggccag	16
•	(210> (211> (212> (213>	596	
•	<220> <221> <222> <223>	(432)(432)	
	<220> <221> <222> <223>	(436)(436)	
	<220> <221> <222> <223>		
	<400> tcgtta	28 aacta ttaggogaat tgggocotot aggtogaoga gttotoagoa gagoogoogt	60
	ctagas	gecee geceteegg gecacegteg gacetagaat agttactega ggtetetegt	120
	cggac	ctaga gtttttcgta tgtttgtcat cgtcggacct aggtccgacg gtccattcct	180
	gagag	totot ctaggtooga ogagagaga aggatootto tgtotagaco otgacgoogg	240
	aaccg	caccg teggacetag gteegacgga aaagcagett cetecactet aggteegacg	300
	gtgtg	tgtgt gtgtgcgtgt tctagagact ggttcagatc aaaagtcgtc ggacctaggt	360
	ccgac	ggggc tggtgagatg gctcagtcta gatgcatgct cgagcggccg ccagtgtgat	420
	ggata	totgo cnaatnocag cacacoggog ogogonacca gtggatocga gcccggtacc	480
	aagct	tgatg catacetega gtatectata etgteaceta aatagettgg ggtaateatg	540
	gtcat	agctg teteetgtgt gaaattgtta teegeteaaa atteecaaca acatag	596

	29 12 DNA Artificial	·	
<220> <223>	linker		
<220> <221> <222> <223>	<pre>misc_feature (1)(1) "n" is any nucleotide</pre>		
<400> nctagg	29 tccg ac		12
<210> <211> <212> <213>	30 20 DNA Artificial		
<220> <223>	linker		
<220> <221> <222> <223>	misc_feature (1)(1) "n" is any nucleotide		
<220> <221> <222> <223>	misc_feature (20)(20) "n" is any nucleotide		
<400> ngttgg	30 acct aggtccaacn		20
<210> <211> <212> <213>	31 13 DNA Artificial		
<220> <223>	linker		
<400> tctagg	31 rtccg acg	·	13
<210> <211> <212> <213>	DNA		
<220> <223>	linker		
<400> cctagg	32 stccg acg		13
<210> <211>	33 20		

## WO 03/106672

<212> <213>	DNA Artificial	
<220> <223>	tagl	
<400> gtggcc	33 cggg agggcggggc	20
	34 19 DNA Artificial	
<220> <223>	tag2	
<400> agagac	34 ctcg agtaactat	19
<210> <211> <212> <213>		
<220> <223>	tag3	
<400> atgaca	35 aaca tacgaaaaac	20
<210> <211> <212> <213>		
<220> <223>	tag4	
<400> gtccat	36 tect gagagtete	19
<210> <211> <212> <213>	37 20 DNA Artificial	
<220> <223>	tag5	
<400> agagag	37 agag gatccttctg	20
<210> <211> <212> <213>	38 20 DNA Artificial	
<220> <223>	tag6	
<400> gtgcgg	38 ttcc ggcgtcaggg	20

## PCT/JP03/07514

<210> <211> <212> <213>	39 19 DNA Artificial	
<220> <223>	tag7	
<400> gaaaag	39 geage tteeteeae	19
<210> <211> <212> <213>	20	
<220> <223>	tag8	
<400> gtgtgt	40 gtgt gtgtgcgtgt	20
<210> <211> <212> <213>	20	
<220> <223>	tag9	
<400> actttt	41 gatc tgaaccagtc	20
<210> <211> <212> <213>	20	
<220> <223>	tag10	
<400> gggctg	42 gtga gatggctcag	20
<210> <211> <212> <213>	19	
<220> <223>	tagl	
<400> gtacct	43 ccctc gcatcccgc	19
<210> <211> <212> <213>	20	
<220> <223>	tag2	
<400> gtggtg	44 rtgcg tgtcgaaggt	20

<212>	20	
<220> <223>	tag3	
<400> atggac	45 cgag ggccccagcc	20
<211>	46 19 DNA Artificial	
<220> <223>	tag4	
<400> cggatc	46 eggt gggtcggac	19
<210> <211> <212> <213>	47 19 DNA Artificial	
<220> <223>	tag5	
<400> agaggt	47 tegea geagttegt	19
<210> <211> <212> <213>	20	
<220> <223>	tag6	
<400> tctccg	48 ggagc cggcgctgtg	20
<210> <211> <212> <213>	19	
<220> <223>	tag7	
<400> agacti	49 ttgca ggctccgag	19
<210> <211> <212> <213>	20 DNA	
<220> <223>	tag8	

<400> ggagct	50 geog cagegeegga	20
<210> <211> <212> <213>		
<220> <223>	tag9 .	
<400> acaccg	51 tcgg acctggtcgc	20
<210> <211> <212> <213>	52 20 DNA Artificial	
<220> <223>	tag10	
<400> agacgt	52 tctc gcccagagtc	20
<210> <211> <212> <213>	53 20 DNA Artificial	
<220> <223>	tag11	
<400> gccgtt	53 cctt gcttgctgga	20
<210> <211> <212> <213>	20	
<220> <223>	tag12	
<400> gggttg	54 ggga tttagctcag	20
<210> <211> <212> <213>	55 19 DNA Artificial	
<220> <223>	tag13	
<400> gagtaa	55 ctat aacggteet	19
<210> <211> <212> <213>	56 19 DNA Artificial	
<220>		

<223>	tag14	
<400> gattcc	56 gcct ggagctcgc	19
<210> <211> <212> <213>	57 18 DNA Artificial	
<220> <223>	tag15	
<400> agggac	57 cgct gcggtccg	18
<210> <211> <212> <213>	58 18 DNA Artificial	
<220> <223>	zzb21106i09t3 junk	
<400> cattag	58 ggga ttgggccc	18
<210> <211> <212> <213>	59 28 DNA Artificial	
<220> <223>	zzb21106i09t3 junk	
<400> acccgg	59 gggg cgggactaac cgtcggac	· 28
<210> <211> <212> <213>	60 20 DNA Artificial	
<220> <223>	tagl	
<400> gacgcg	60 gaag gegeggegge	20
<210> <211> <212> <213>	61 21 DNA Artificial	
<220> <223>	tag2	
<400> ggaggg	61 cggg cggcggccct c	21
<210> <211> <212> <213>	62 19 DNA Artificial	

WO 03/106672 PCT/JP03/07514

<220> <223>	tag3	
<400> caaaaaa	62 aaaa aaaaaaact	19
<210>	63	
<211> <212> <213>	20 DNA Artificial	
<220> <223>	tag4	
	63 tgct cgctctgccc	20
<210> <211> <212> <213>	64 19 DNA Artificial	
<220> <223>	tag5	
<400> acttct	64 gatt ctgacagac	19
<210> <211> <212> <213>	65 19 DNA Artificial	
<220> <223>	tag6	
<400> acagtg	65 gcgt ctgcaaagc	19
<210> <211> <212> <213>	66 20 DNA Artificial	
<220> <223>	tag7	
<400> ggaaag	66 tcca ggtggacttt	20
<210> <211> <212> <213>	67 19 DNA Artificial	
<220> <223>	tag8	
<400> gccgcc	67 gagg ccgcgcagg	19
<210> <211>	68 19	

misc\_feature
(1)..(1)
"n" is any nucleotide

PCT/JP03/07514

<pre>&lt;400&gt; 73 nccatggaac agccacact</pre>	19
<210> 74 <211> 26 <212> DNA <213> Artificial	
<220> <223> junk1	·
<400> 74 tgataaggca atggcctcta atgctg	26
<210> 75 <211> 16 <212> DNA <213> Artificial	
<220> <223> tag	
<400> 75 acctccctcc gcggag	16
<210> 76 <211> 2745 <212> DNA <213> Homo sapiens	
<220> <221> CDS <222> (842)(2017) <223>	
<400> 76 acctcctcc gcggagcagc cagacagcga gggccccggc cgggggcagg ggggacgcc	c 60
cgtccggggc acccccccg gctctgagcc gcccgcgggg ccggcctcgg cccggagcg	g 120
aggaaggagt cgccgaggag cagcctgagg ccccagagtc tgagacgagc cgccgccgc	c 180
cocgccactg oggggaggag ggggaggagga gggacgagct ggtcgggag	a 240
agaggaaaaa aacttttgag acttttccgt tgccgctggg agccggaggc gcggggacc	t 300
cttggcgcga cgctgccccg cgaggaggca ggacttgggg accccagacc gcctccctt	t 360
geogeogggg acgettgete cetecetgee cectacaegg egteceteag gegeeecea	t 420
teeggaccag ecetegggag tegeegacce ggeeteegg aaagaetttt eeceagace	t 480
eggegeace ecetgeacge egeetteate eceggeetgt etectgagee ecegegeat	c 540
ctagaccett teteeteeag gagaeggate teteteegae etgeeacaga teceetatt	c 600
aagaccaccc accttctggt accagatcgc gcccatctag gttatttccg tgggatact	g 660
agacaccccc ggtccaagcc tcccctccac cactgcgccc ttctccctga ggagcctca	g 720
ctttccctcg aggecetect accttttgcc gggagacccc cagecectge aggggeggg	g 780
cctccccacc acaccagccc tgttcgcgct ctcggcagtg ccgggggggg ccgcctccc	c 840
c atg ccg ccc tcc ggg ctg cgg ctg ctg ccg ctg ct	s 889

]	Met 1	Pro	Pro	Ser	Gly 1 5	Leu A	Arg 1	Leu 1	Leu	Pro 1	Leu	Leu	Leu	Pro	Leu Le 15	u
tg: Tr:	g ct p Le	a ct u Le	g gt u Va 20	T PG	g aca u Thi	g cci	t ggo o Gl	c ccay Pro 25	g cc; o Pr	g gc	c gc a Al	g gg a Gl	a ct y Le 30	a tc u Se	c acc r Thr	937
tg:	c aaq s Ly:	g ac s Th 35	1 11	c ga e As	c ata p Me	g gag t Glu	cta Let 40	g gta	g aaq l Ly:	g cgg	g aa g Ly	g cg s Ar 45	c ato	c ga e Gl	g gcc u Ala	985
ato Ile	c cge Arg 50	c gg g Gl;	c can	g ato n Ilo	c cta e Lei	tco Ser 55	aag Lys	cta Lei	g cgg	s cto s Leu	gc Al 60	c ag a Se	c cco r Pro	c cc	g agc o Ser	1033
65	1 41)	/ <b>G</b> II	u va.	i Pro	70	GI3	Pro	) Let	ı Pro	75	ı Ala	a Va.	l Let	ı Ala	c ctg Leu 80	1081
				00					90					95	g gag Glu	1129
	ulu		100	)	r voh	ı ıyr	ıyr	105	Lys	610	va.	ITN	110	Ya.	cta Leu	1177
ric c	141	115	, 1111	urs	ASII	. GIU	120	lyr	' Asp	) Lys	Phe	125	Gln	Ser	aca Thr	1225
1113	130	116	; 131	riet	rue	135	ASII	ınr	Ser	Glu	140	Arg	Glu	Ala	gta Val	1273
145	uru	110	, , 4	neu	150	ser	wirk	Ala	aru	155	Arg	Leu	Leu	Arg	agg Arg 160	1321
				100					170					175	agc Ser	1369
11311	non	PCI	180	шg	ŢŢŢ	rea	ser	185	Arg	Leu	Leu	Ala	190	Ser	gac Asp	1417
	110	195	11.5	Leu	tct Ser	rne	200	vai	ınr	GIÀ	Val	205	Arg	Gin	Trp	1465
	2.0				gaa Glu	213					220					1513
225	0,3	пор	DCI	we.	gat Asp 230	nsii	Inr	Leu	GID	235	Asp	He	Asn	Gly	Phe 240	1561
1111	1111	ulj	шŞ	245	ggt Gly	ASP	Leu	ATA	1nr 250	He	His	Gly	Met	Asn 255	Arg	1609
. 10	ı nc		260	Deu	atg Met	via	ш	265	rea	ara	Arg	Ala	G1n 270	HIS	Leu	1657
caa Gln	DCI	Ser 275	cgg Arg	cac His	cgc Arg	w.g	gcc Ala 280	ctg Leu	gac Asp	acc Thr	aac Asn	tat Tyr 285	tgc Cys	ttc Phe	agc Ser	1705

tcc acg gag aag aac tgc tgc gtg cgg cag ctg tac att gac ttc cgc Ser Thr Glu Lys Asn Cys Cys Val Arg Gin Leu Tyr Ile Asp Phe Arg 290 295 300	1753
aag gac ctc ggc tgg aag tgg atc cac gag ccc aag ggc tac cat gcc Lys Asp Leu Gly Trp Lys Trp Ile His Glu Pro Lys Gly Tyr His Ala 305 310 315 320	1801
aac ttc tgc ctc ggg ccc tgc ccc tac att tgg agc ctg gac acg cag Asn Phe Cys Leu Gly Pro Cys Pro Tyr Ile Trp Ser Leu Asp Thr Gln 325 330 335	1849
tac agc aag gtc ctg gcc ctg tac aac cag cat aac ccg ggc gcc tcg Tyr Ser Lys Val Leu Ala Leu Tyr Asn Gln His Asn Pro Gly Ala Ser 340 345 350	1897
gcg gcg ccg tgc tgc gtg ccg cag gcg ctg gag ccg ctg ccc atc gtg Ala Ala Pro Cys Cys Val Pro Gln Ala Leu Glu Pro Leu Pro Ile Val 355 360 365	1945
tac tac gtg ggc cgc aag ccc aag gtg gag cag ctg tcc aac atg atc Tyr Tyr Val Gly Arg Lys Pro Lys Val Glu Gln Leu Ser Asn Met Ile 370 375 380	1993
gtg cgc tcc tgc aag tgc agc tga ggtcccgccc cgccccgccc	2047
ggcccggccc caccccgccc cgcccccgct gccttgccca tggggggctgt atttaaggac	2107
acceteccc aagcccacct ggggccccat taaagatgga gagaggactg cggatctctg	2167
tgtcattggg cgcctgcctg gggtctccat ccctgacgtt cccccactcc cactccctct	2227
ctctccctct ctgcctcctc ctgcctgtct gcactattcc tttgcccggc atcaaggcac	2287
aggggaccag tggggaacac tactgtagtt agatctattt attgagcacc ttgggcactg	2347
ttgaagtgcc ttacattaat gaactcattc agtcaccata gcaacactct gagatggcag	2407
ggactctgat aacacccatt ttaaaggttg aggaaacaag cccagagagg ttaagggagg	2467
agtteetgee caccaggaac etgetttagt gggggatagt gaagaagaca ataaaagata	2527
gtagttcagg ccaggcgggg tgctcacgcc tgtaatccta gcacttttgg gaggcagaga	2587
tgggaggata cttgaatcca ggcatttgag accagcctgg gtaacatagt gagaccctat	2647
ctctacaaaa cacttttaaa aaatgtacac ctgtggtccc agctactctg gaggctaagg	2707
tgggaggatc acttgatcct gggaggtcaa ggctgcag	2745

<210> 77

<211> 391 <212> PRT

<213> Homo sapiens

<400> 77

Trp Leu Leu Val Leu Thr Pro Gly Pro Pro Ala Ala Gly Leu Ser Thr  $\frac{1}{20}$ 

Cys Lys Thr Ile Asp Met Glu Leu Val Lys Arg Lys Arg Ile Glu Ala

35

40

45

Ile Arg Gly Gln Ile Leu Ser Lys Leu Arg Leu Ala Ser Pro Pro Ser 50 60

Gln Gly Glu Val Pro Pro Gly Pro Leu Pro Glu Ala Val Leu Ala Leu 65 70 75 80

Tyr Asn Ser Thr Arg Asp Arg Val Ala Gly Glu Ser Ala Glu Pro Glu 85 90 95

Pro Glu Pro Glu Ala Asp Tyr Tyr Ala Lys Glu Val Thr Arg Val Leu 100 105 110

Met Val Glu Thr His Asn Glu Ile Tyr Asp Lys Phe Lys Gln Ser Thr

His Ser Ile Tyr Met Phe Phe Asn Thr Ser Glu Leu Arg Glu Ala Val 130 135 140

Pro Glu Pro Val Leu Leu Ser Arg Ala Glu Leu Arg Leu Leu Arg 145 150 160

Leu Lys Leu Lys Val Glu Gln His Val Glu Leu Tyr Gln Lys Tyr Ser 165 170 175

Asn Asn Ser Trp Arg Tyr Leu Ser Asn Arg Leu Leu Ala Pro Ser Asp 180 180 190

Ser Pro Glu Trp Leu Ser Phe Asp Val Thr Gly Val Val Arg Gln Trp 195 200 205

Leu Ser Arg Gly Gly Glu Ile Glu Gly Phe Arg Leu Ser Ala His Cys 210 215

Ser Cys Asp Ser Arg Asp Asn Thr Leu Gln Val Asp Ile Asn Gly Phe 225 230 235 240

Thr Thr Gly Arg Arg Gly Asp Leu Ala Thr Ile His Gly Met Asn Arg 245 255

Pro Phe Leu Leu Met Ala Thr Pro Leu Glu Arg Ala Gln His Leu 260 270

Gln Ser Ser Arg His Arg Arg Ala Leu Asp Thr Asn Tyr Cys Phe Ser 275 280

Ser Thr Glu Lys Asn Cys Cys Val Arg Gln Leu Tyr Ile Asp Phe Arg 290 295 300

Lys Asp Leu Gly Trp Lys Trp Ile His Glu Pro Lys Gly Tyr His Ala 305 310 315 320 WO 03/106672 PCT/JP03/07514

Tyr Ser Lys Val Leu Ala Leu Tyr Asn Gln His Asn Pro Gly Ala Ser 340

Ala Ala Pro Cys Cys Val Pro Gln Ala Leu Glu Pro Leu Pro Ile Val 355

Tyr Tyr Val Gly Arg Lys Pro Lys Val Glu Gln Leu Ser Asn Met Ile 370 375 380

Val Arg Ser Cys Lys Cys Ser 385